Design & Development of Real-Time HCI for Hand-free PC Operations using Image Processing

Pooja R. Kapse[1], Prof. Yogesh S. Patil[2], Prof. Dinesh D. Patil[3]

[1] *M.E. Student, Dept. of Computer Science and Engineering, SSGBCOET, Bhusawal, Maharashtra, India*
[2] *Assistant Professor, Computer Science and Engineering, SSGBCOET, Bhusawal, Maharashtra, India*
[3] *HOD, Computer Science and Engineering, SSGBCOET, Bhusawal, Maharashtra, India*

ABSTRACT

Computers have become an important part of our daily life but some people like handicapped/paralyzed are not able to use the computers because of their disabilities. Face detection has been an important and active research field because it offers many applications, especially in video surveillance, biometrics, or video coding. Facial expressions are used to control the cursor movement in computer applications. The aim of this system is to use human body movement tracking in video input and utilize the tracking is an important kind of solution. In order to facilitate people with disabilities, use of the face recognition and extraction of user eyes location is done. The system uses Viola Jones algorithm which uses Haar-like features, that is, a scalar product between the image and some Haar-like templates. Various experiments were conducted to verify the results that were more accurate than the current solution to camera mouse. It also identifies the problems associated with past developments of the system. However the main purpose is to put forward a modified version of the system.

Keyword: *Eye Detection, Face Detection, Viola Jones Algorithm, Haar-like Features, Mouse Controlling.*

1. INTRODUCTION

Researchers are studying new techniques to ease the human-computer interaction. Generally human-computer interface consists of devices where inputs are acquired and corresponding outputs are generated. The most basic input devices are keyboard, mouse and mice, etc. whereas output devices are monitor screen, loudspeaker and printer, etc. But in case of disabled people all these input devices are useless. Face detection is the process of determining whether a face is present in an image or not. Unlike face recognition-which distinguishes different human faces, face detection only indicates presence of a face in an image. In addition, eye detection traces or extracts the location of user's eye. This method is very useful for interaction with the computer without having any physical connection.

Usually, there are several parts of body can be used for tracking and apperceiving user interface such as: face, mouth, hand, eyes, etc. there are also many kinds of algorithm for these parts such as: navigate mouse using eyes movement information, using nose movement and using nostrils movement, etc. using face as an input device for computer, with the help of web camera. Where the face is detected, its pattern and then using various algorithmic techniques they are computed and changes on face are converted into actions of mouse. Various facial expressions are used to carry out different mouse actions like click, double click, scrolling, and zoom, etc. The movement of cursor is done based on the movement of user's face.

Various examples for face detection system are used for checking whether or not the person is looking directly at the TV. If the person is not directly looking at the TV within some time period (i.e. 15 minutes), the TV's brightness is reduced to save energy. When the person turns back to look at the TV, the TV's brightness can be increased back to original. In addition, if the person looks away for too long (i.e. more than one hour), then the TV will be automatically turned off.

2. LITERATURE SURVEY

For human-computer interface hand-free control has been very important nowadays. To use human body movement tracking for hand-free solution in video input and utilize the tracking is an important kind of solution. A hand-free head mouse control based on mouth tracking system is proposed where firstly transformation of nonlinearly input video frame of human head into YC_bC_r color space is done, then the visible chrominance feature of face in this color space to detect human face region is used [1]. And then for face candidate, use of nearly reversed relationship information between C_b and C_r cluster of face feature is used to detect mouth position.

Perceptual User Interfaces (PUI) is a highly interactive, multimodal interfaces modeled after natural human-to-human interaction, with the goal of enabling people to interact with technology in a similar fashion to how they interact with each other and with the physical world. Vision Based Interfaces (VBI) is a subfield of perceptual interfaces which concentrates on developing visual awareness of people where computer vision algorithms are implemented to locate and identify individuals, track human body motions, model the head and face, track facial features, interpret human motion and actions [2].

A robust, accurate, and low cost real-time solution for the eye and face detection problem is introduced where the method uses two infrared illumination sources to generate bright and dark pupil images, which are combined to robustly detect pupils [3]. Once the pupils are detected, the inter-ocular distance is used to determine the size and position of the bounding box around the face.

The human factors and technical considered that arise in trying to use eye movements as an input medium. The goal was to measure visual line of gaze, that is, the absolute position in space at which the user's eyes are pointed, rather than, for example, the position of the eyeball in space or the relative motion of the eye within the head [4]. A multi-camera vision-based eye tracking method is introduced to robustly locate and track user's eyes as they interact with an application. The enhancements were proposed to various vision-based eye-tracking approaches, which include (a) the use of multiple cameras to estimate head pose and increase coverage of the sensors and (b) the use of probabilistic measures incorporating Fisher's linear discriminant to robustly track the eyes under varying lighting conditions in real-time [5].

The concept of face tracking is deeply focused in [6] [7] where, the theoretical and by experiments conducted with ordinary USB cameras the system shows that, by properly defining nose - as an extremum of the 3D curvature of the nose surface, nose becomes the most robust feature which can be seen for almost any position of the head and which can be tracked very precisely even with low resolution cameras, whereas, the vision-based technology which allows one in such a setup to significantly enhance the perceptual power of the computer. The described techniques for tracking a face using a convex-shape nose feature as well as for face-tracking with two off-the-shelf cameras allow one to track faces robustly and precisely in both 2D and 3D with low resolution cameras [7].

Motivated by the goal of providing a non-contact means of controlling the mouse pointer on a computer system for people with motor difficulties using low-cost, widely available hardware. The required information is derived from video data captured using a web camera mounted below the computer's monitor. A color filter is used to identify skin colored regions.

False positives are eliminated by optionally removing background regions and by applying statistical rules that reliably identify the largest skin-colored region, which is assumed to be the user's face [8].

The implementation of a system such as the proposed one presents several areas of difficulty:

1. Identifying and tracking the head location.
2. Identifying and tracking the location of a facial feature.
3. Being able to process the information in real-time using a moderately priced processor that will be running other applications in the foreground (for example, Microsoft Word).

Hands-free Interface is an assistive technology that is intended mainly for the use of the disabled. It would help them use their voluntary movements, like head movements; to control computers and communicate through customized educational software or expression building programs. People with severe disabilities can also benefit from computer access to partake in recreational activities, use the Internet or play games. The system uses a USB camera to capture the user's face motion. The algorithm tracks the motion accurately to control the cursor, thus providing an alternative to the computer mouse or keyboard [9]. Also, a face tracking technique was introduced that solves the aforementioned problems and can be used for various real time applications.

Detecting faces in images with complex backgrounds is a difficult task. The approach, which obtains state of the art results, is based on a new neural network model: the Constrained Generative Model (CGM). Generative, since the goal of the learning process is to evaluate the probability that the model has generated the input data, and constrained since some counterexamples are used to increase the quality of the estimation performed by the model.

To detect side view faces and to decrease the number of false alarms, a conditional mixture of networks is used. To decrease the computational time cost, a fast search algorithm was proposed. The level of performance reached, in terms of detection accuracy and processing time, allows applying this detector to a real world application: the indexation of images and videos [10].

Humans detect and interpret faces and facial expressions in a scene with little or no effort. Still, development of an automated system that accomplishes this task is rather difficult. There are several related problems: detection of an image segment as a face, extraction of the facial expression information, and classification of the expression (e.g., in emotion categories). A system that performs these operations accurately and in real time would form a big step in achieving a human-like interaction between man and machine. The capability of the human visual system with respect to these problems also considered. It is meant to serve as an ultimate goal and a guide for determining recommendations for development of an automatic facial expression analyzer [11].

A face detection algorithm for color images in the presence of varying lighting conditions as well as complex backgrounds. Our method detects skin regions over the entire image, and then generates face candidates based on the spatial arrangement of these skin patches [12]. The algorithm constructs eye, mouth, and boundary maps for verifying each face candidate. Experiments were conducted and results demonstrate successful detection over a wide variety of facial variations in color, position, scale, rotation, pose, and expression from several photo collections. The use of color information can simplify face localization in complex environments. An overview of face detection algorithm which contains two major modules: (i) face localization for finding face candidates; and (ii) facial feature detection for verifying detected face candidates.

A synthetic exemplar based framework for face recognition with variant pose and illumination as proposed. Our purpose is to construct a face recognition system only according to one single frontal face image of each person for recognition. The framework consists of three main parts. First, a deformation based 3D face modeling technique is introduced to create an individual 3Dface model from a single frontal face image of a person with a generic 3D face model. Then, the virtual faces for recognition at various lightings and views are synthesized finally, an Eigenfaces based classifier is constructed where the virtual faces synthesized are used as training exemplars. The experimental results showed that the 3D face modeling technique is efficient and the synthetic face exemplars can significantly improve the accuracy of face recognition with variant pose and illumination [13].

Camera position information from2D face image is very important for that make the virtual 3D face model synchronize to the real face at view point, and it is also very important for any other uses such as human-computer interface (face mouth), automatic camera control, etc. A algorithm is presented to detect human face region and mouth, based on special color features of face and mouth in YC_bC_r color space. The algorithm constructed a mouth feature image based on C_b and C_r values, and use pattern method to detect the mouth position. And then the geometrical relationship is used between mouth position information and face side boundary information to determine the camera position. The experiment results demonstrated the validity of the algorithm and the Correct Determination Rate is accredited for applying it into practice [14].

3. WORKING OF THE SYSTEM

The block diagram of the system is shown in the figure 1. As shown in the figure, the proposed system consists of 3 main parts:

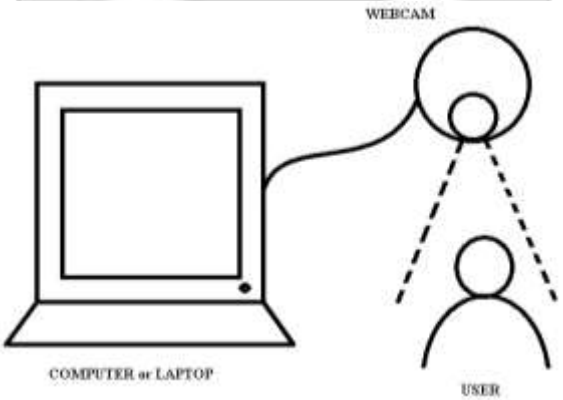

Figure 3.1: Block Diagram of Proposed System.

i. Computer or Laptop:

The computer or laptop is required to perform the specified task. The mouse controlling and triggering of pre-defined PC operations can be operated using a PC or laptop. A computer is a device that can be instructed to carry out an arbitrary set of arithmetic or logical operations automatically. The ability of computers to follow a sequence of operations, called a program, make computers very flexible and useful. Such computers are used as control systems for a very wide variety of industrial and consumer devices. This includes simple special purpose devices like microwave ovens and remote controls, factory devices such as industrial robots and computer assisted design, but also in general purpose devices like personal computers and mobile devices such as smartphones. The Internet is run on computers and it connects millions of other computers.

ii. User:

The User is required and is the most important part of the proposed system because the mouse will be operated through the actions of the user by making use of the camera. A system user is a person who interacts with a system, typically through an interface, to extract some functional benefit. User-centered design, often associated with human–computer interaction, considers a wide range of generic systems. System user also defines the behavior of the system operations and how the audience (end-user) would interact with the system using pre-designed triggers such as buttons/mouse/keyboard.

iii. Webcam:

The webcam is the third entity of the setup. A webcam is a video camera that feeds or streams its image in real time to or through a computer to computer network. When "captured" by the computer, the video stream may be saved, viewed or sent on to other networks via systems such as the internet, and email as an attachment. When sent to a remote location, the video stream may be saved, viewed or on sent there. Unlike an IP camera (which connects using Ethernet or Wi-Fi), a webcam is generally connected by a USB cable, or similar cable, or built into computer hardware, such as laptops. The term "webcam" (a clipped compound) may also be used in its original sense of a video camera connected to the Web continuously for an indefinite time, rather than for a particular session, generally supplying a view for anyone who visits its web page over the Internet. Some of them, for example, those used as online traffic cameras, are expensive, rugged professional video cameras.

The system works on the flow given below, which can be seen in figure 3.2.

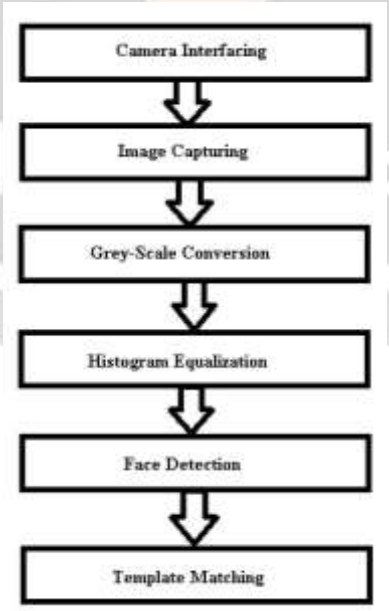

Figure 3.2: Method of the system.

i. Camera Interfacing:

The webcam is needed to be connected to the computer in order to obtain the live feed for the extraction of the images of the user. This image will be used in the entire process.

ii. Image Capturing:

The webcam will obtain a live video through which images of the user will be captured. These images will be processed further to form frames.

iii. Grey-Scale Conversion

A grayscale or greyscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Grayscale images are distinct from one-bit bi-tonal black-and-white images, which in the context of computer imaging are images with only two colors, black and white (also called bi-level or binary images). Grayscale images have many shades of gray in between. Figure 3.3 shows a sample of grayscale image.

Figure 3.3: Sample of Grayscale Image.

iv. Histogram Equalization:

Histogram equalization is a method in image processing of contrast adjustment using the image's histogram. It can also be used on color images by applying the same method separately to the Red, Green and Blue components of the RGB color values of the image. However, applying the same method on the Red, Green, and Blue components of an RGB image may yield dramatic changes in the image's color balance since the relative distributions of the color channels change as a result of applying the algorithm. However, if the image is first converted to another color space, Lab color space, or HSL/HSV color space in particular, then the algorithm can be applied to the luminance or value channel without resulting in changes to the hue and saturation of the image. There are several histogram equalization methods in 3D space. Trahanias and Venetsanopoulos applied histogram equalization in 3D color space. However, it results in "whitening" where the probability of bright pixels is higher than that of dark ones. Histogram equalization example can be seen in figure 3.4.

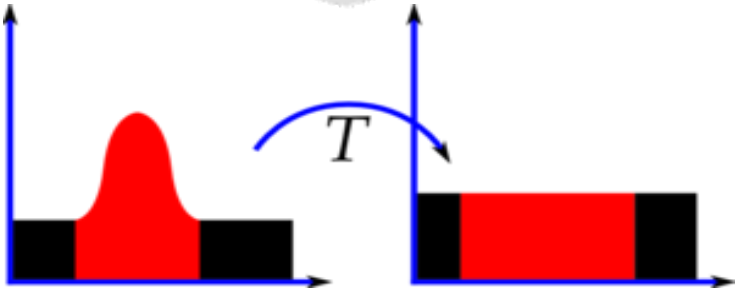

Figure 3.4: Sample of Histogram Equalization.

v. Face Detection

Face detection is a computer technology being used in a variety of applications that identifies human faces in digital images. Face detection also refers to the psychological process by which humans locate and attend to faces in a visual scene. Face detection process is shown in figure 3.5.

Figure 3.5: Face Detection Process

Face Detection can be applied in the following fields:

Face Recognition: Face detection is used in biometrics, often as a part of (or together with) a facial recognition system. It is also used in video surveillance, human computer interface and image database management.

Photography: Some recent digital cameras use face detection for autofocus. Face detection is also useful for selecting regions of interest in photo slideshows that use a pan-and-scale Ken Burns effect. Modern appliances also use smile detection to take a photograph at an appropriate time.

Marketing: Face detection is gaining the interest of marketers. A webcam can be integrated into a television and detect any face that walks by. The system then calculates the race, gender, and age range of the face. Once the information is collected, a series of advertisements can be played that is specific toward the detected race/gender/age.

vi. Template Matching

Template matching is a technique in digital image processing for finding small parts of an image which match a template image. It can be used in manufacturing as a part of quality control, a way to navigate a mobile robot, or as a way to detect edges in images.

4. ADVANTAGES

There are many advantages as listed below:

1. It is helpful for usability studies to understand user's interaction with their environments.
2. Disabled people can interact with computer easily.
3. It is a very flexible method as more features of face gestures are provided.

5. APPLICATIONS

There are number of applications of the developed system, some are given below:

1. Hand-free computing is useful to both able and disabled people. This means that no physical connection is present between the user and the system.
2. It can be used in development video games and graphics.
3. It provides new and more effective methods of computer-human interaction.

6. CONCLUSION

The system can be useful in many different ways. It makes use of the RGB color space. It is implemented in OpenCV as it provides Haar classifier. It is based on Computer Vision technology. The system to detect eye blink in real-time with low cost hardware setup and processing it for movement of mouse and track human faces was achieved. The work overcomes the current issue in face movement controlled mouse after going through various systems currently available. The system will be helpful for the disabled people without needing the costly devices that are planted in human body for interaction with the computer.

7. REFERENCES

[1] Zhang Naizhong, Wen Jing, Wang Jun "Hand-Free Head Mouse Control Based on Mouth Tracking," IEEE International Conference on Computer Science & Education, pp. 707-713, 2015.

[2] M. Turk, C. Hu, R. Feris, F. Lashkari and A. Beall, "TLA based face tracking," International Conference on Vision Interface, pp.229-235, 2002.

[3] C. Morimoto, D. Koons, A. Amir, and M. Flickner, "Realtime detection of eyes and faces, In Workshop on Perceptural User Interfaces," San Fransisco, 1998.

[4] R. J. Jacob, "What you look at is what you get: Eye movement based interaction techniques," In Human Factors in Computing Systems, pp.11-18, 1990.

[5] R. Ruddarraju and etal, "Perceptual user interfaces using vision based eye tracking," In the 5th international conference on Multimodal interface, pp. 227-233.s, 2003.

[6] D. Gorodnichy. "On importance of nose for face tracking, Automatic Face and Gesture Recognition," Fifth IEEE International Conference, pp.181-186, 2002.

[7] D. Gorodnichy, S. Malik, and G. Roth, "Noose' use your nose as a mouse' - a new technology for hands-free games and interfaces," Intern. Conf. on Vision Interface, pp.354-361, 2002.

[8] V.Chauhan and T.Morris, "Face and feature tracking for cursor control," In 12th Scandinavian Conference on Image Analysis, Bergen, 2001.

[9] L. El-Afifi, M. Karaki, and J. Korban, "'hand-free interface'- a fast and accurate tracking procecdure for real time human computer interaction," In FEA Student Conference, American University of Beirut, 2004.

[10] R.Feraud, O.J.Bernier, J.-E. Viallet, and M. Collobert, "A Fast and Accurate Face Detection Based on Neural Network," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.23, no. 1, pp.42-53, 2001.

[11] M. Pantic and L.J.M. Rothkrantz, "Automatic Analysis of Facial Expressions: The State of the Art," IEEE Trans, Pattern Analysis and Machine Intelligence, Vol.22, no. 12, pp. 1424-1445, 1996.

[12] Rein-Lien Hsu and Mohamed Abdel-Mottaleb, "Face Detection in Color Images," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24, pp. 696 - 706, 2002.

[13] Yuankui Hu, Ying Zheng and Zengfu Wang, "Reconstruction of 3D Face from a Single 2D Image for Face Recognition," Joint IEEE International Workshop on pp.7-222, 15-16 Oct. 2005.

[14] Li Wei, E.J. Lee, "Human Head Mouse System Based on Facial Gesture Recognition," Journal of Korea Multimedia Society, Vol. 10, No. 12, pp. 1591-1600, Dec 2007.